Optimal Sampling for State Change Detection with Application to the Control of Sleep Mode

Amar P. Azad*, Sara Alouf*, Eitan Altman*, Vivek Borkar† and Georgios Paschos+
* Maestro group, INRIA, 2004 Route des Lucioles, F-06902 Sophia Antipolis
†School of Technology, TIFR, Mumbai
+ CERTH, Thessaloniki, Greece
{aazad,salouf,altman}@sophia.inria.fr, borkar@tifr.res.in, gpasxos@ee.upatras.gr

Abstract

This work considers systems with inactivity periods of unknown duration. We study the question of scheduling "waking up" instants in which a server can check whether the inactivity period is over. There is a cost proportional to the delay from the moment the inactivity period ends until the server discovers it, a (small) running cost while the server is away and also a cost for waking up. As an application to the problem, we consider the energy management in WiMax where inactive mobiles reduce their energy consumption by entering a sleep mode. Various standards exist which impose specific waking-up scheduling policies at wireless devices. We check these and identify optimal policies under various statistical assumptions. We show that periodic fixed vacation durations are optimal and derive the optimal period. We show that this structure does not hold for other inactivity distributions but manage to obtain some suboptimal solutions which perform strictly better than the periodic ones. We finally obtain structural properties for optimal policies for the case of arbitrary distribution of inactivity periods.

1 Introduction

Mobile terminals using contemporary radios can benefit greatly by shutting off the transceiver whenever there is no scheduled activity. Nevertheless, if the attention of the mobile is suddenly required, the mobile will be shut off and therefore unavailable. The longer the shut off (vacation) periods, the longer the expected response delay. Therefore, one can identify the inherent tradeoff of energy management: increase vacation length to improve energy saving or decrease vacation length to reduce delays.

Past approaches (see references in [1]) have considered incoming/outgoing traffic, the effect of setup time, or even the queueing implications in the analysis. Concerning the arrival process, it has been assumed to be Poisson, having a hyper-Erlang distribution or a hyper-exponential distribution. In all cases, it does not depend on the energy management scheme. As for delay, it is the average packet delay in the system that is considered.

Recent works [8, 12, 7] focus on heuristic adaptive algorithms whose goal is to control the vacation length according to the incoming arrival process. The work [10] derives an optimal sleep policy using average cost structure for a given number of consecutive sleep durations.

Our work departs from the existing models in two aspects. First, rather than an exogenous independent arrival process, we have in mind elastic arrival processes in which (i) a "think time" or "off time" begins when the activity of the server ends, and (ii) the duration of the "on time" does not depend on the wake up delay, defined as the time that elapses between the instant a request is issued and the instant at which the request service actually begins. Both assumptions are appropriate to interactive applications such as web browsing. As a result, the measure for delay is taken to be the wake up delay.

Our objective is to optimize the vacation duration in order to achieve the desired balance between delay and energy saving. We shall investigate in this paper optimal energy management systems under one of the following assumptions on the off time distribution: (i) *Exponential distribution*; (ii) *Hyper-exponential distribution*; (iii) *General distribution*. The motivation behind the hyper-exponential distribution assumption comes from works that provide evidence of heavy-tailed off time distributions on the Internet [11] and of Pareto type distribution on the World Wide Web [3]. Furthermore, it is well-known that heavy-tailed distributed random variables (rvs) can be well approximated by hyper-exponential distributions [5].

Our contributions are as follows: (1) Our problem formulation allows us to minimize the weighted sum of the two costs, which is essentially obtaining the optimal tradeoff of delay against energy saving. We use dynamic programming (DP) which allows to obtain the optimal vacation size at each wake up instant. (2) For exponential off times, we show that the constant vacation policy is optimal and we derive it. (3) For hyper-exponential off times, we derive interesting structural properties. We show that the optimal control is bounded. Asymptotically, the optimal policy converges to the constant policy corresponding to the smallest rate phase, irrespective of the initial state. This policy can be computed numerically using value iteration. (4) For any general off time distribution, we show that the optimal control is bounded. (5) We propose suboptimal policies using policy iteration which perform strictly better than optimal "homogeneous" policies and are simpler to compute. We show numerically the performance of such suboptimal solutions using one stage and two stage policy iteration. (6) We compare the proposed policies with that of the IEEE 802.16e standard [6] under various statistical assumptions.

In the rest of the paper, Sect. 2 outlines our system model and introduces the cost function. Section 3 introduces DP and derives the optimal sleep control and relevant characteristics for hyper-exponential off times. Section 4 tackles the problem of finding the optimal policy under the worst case process of arrivals. Numerical results and a comparative study of the different (sub)optimal strategies and of the IEEE 802.16e standard are reported in Sect. 5. Section 6 concludes the paper. Due to space limitations, most proofs are omitted from the paper, but can be found in [1].

2 System Model

We consider a system with repeated vacations. As long as there are no customers, the server goes on vacation. We are interested in finding the optimal policy, so that at any start of vacation, the length of this vacation is optimal. This system models a mobile device that turns off its radio antenna while inactive to save energy. A vacation is then the time during which the mobile is sleeping. At the end of a vacation, the mobile needs to turn on the radio to check for packets.

Let X denote the number of vacations in an idle period. X is a discrete random variable (rv) taking values in $\mathbb{N}^*$. The duration of the kth vacation is a rv denoted B_k, for $k \in \mathbb{N}^*$. For analytical tractability, we consider vacations $\{B_k\}_{k\in\mathbb{N}^*}$ that are mutually independent rvs. The time at the end of the kth sleep interval is a rv denoted T_k, for $k \in \mathbb{N}^*$. We denote T_0 as the time at the beginning of the first vacation; by convention $T_0 = 0$. We naturally have $T_k = T_{k-1} + B_k = \sum_{i=1}^{k} B_i$. Observe that a generic idle ends at time T_X.

We will be using the following notation $\mathcal{Y}(s) := \mathbb{E}[\exp(-sY)]$ to denote the Laplace-Stieltjes transform of a generic rv Y evaluated at s. Hence, we can readily write $\mathcal{T}_k(s) = \prod_{i=1}^{k} \mathcal{B}_i(s)$.

Let τ denote the time length between the start of the first vacation and the arrival of a customer; this time is referred to as the "off time". Since a generic idle period ends at time T_X, the service of the first customer to arrive during the idle period will be delayed for $T_X - \tau$ units of time.

τ is a rv whose probability density function is $f_\tau(t), t \geq 0$. We will be assuming that τ is hyper-exponentially distributed with n phases and parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$. In other words, we have

$$f_\tau(t) = \sum_{i=1}^{n} q_i \lambda_i \exp(-\lambda_i t), \qquad \sum_{i=1}^{n} q_i = 1. \tag{1}$$

Given its definition, the off time τ is also the conditional residual inter-arrival time. Observe that when $n = 1$, τ will be exponentially distributed with parameter $\lambda = \lambda_1$, which, thanks to the memoryless property of this distribution, is equivalent to having a Poisson arrival process with rate λ.

The energy consumed by a mobile while *listening* to the channel and checking for customers is denoted E_L. This is actually a penalty paid at the end of each vacation. The *power* consumed by a mobile in a sleep state is denoted P_S. The energy consumed by a mobile during vacation B_k is then equal to $E_L + P_S B_k$, and that consumed during a generic idle period is equal to $E_L X + P_S T_X$.

We are interested in minimizing the cost of the power save mode, which is seen as a weighted sum of the energy consumed during the power save mode and the *extra* delay incurred on the traffic by a sleeping mobile. Let V be this cost; it is written as follows

$$\begin{aligned} V &:= \mathbb{E}\left[\bar{\epsilon}\,(T_X - \tau) + \epsilon\,(E_L X + P_S T_X)\right] & (2)\\ &= -\bar{\epsilon}\,\mathbb{E}[\tau] + \epsilon E_L \mathbb{E}[X] + \eta \mathbb{E}[T_X] & (3)\end{aligned}$$

where ϵ is a *normalized weight* that takes value between 0 and 1, $\bar{\epsilon} = 1 - \epsilon$, $\eta := \bar{\epsilon} + \epsilon P_S$. The derivation of the elements of (3) when τ is hyper-exponentially distributed is straightforward. We derive

$$\begin{aligned} P(X = k) &= P(\tau > T_{k-1}) - P(\tau > T_k)\\ &= \sum_{i=1}^{n} q_i \mathcal{T}_{k-1}(\lambda_i)\,(1 - \mathcal{B}_k(\lambda_i))\,;\end{aligned}$$

$$\mathbb{E}[\tau] = \sum_{i=1}^{n} q_i/\lambda_i; \qquad \mathbb{E}[X] = \sum_{k=0}^{\infty}\sum_{i=1}^{n} q_i \mathcal{T}_k(\lambda_i); \tag{4}$$

$$\mathbb{E}[T_X] = \sum_{k=0}^{\infty}\sum_{i=1}^{n} q_i \mathcal{T}_k(\lambda_i)\mathbb{E}[B_{k+1}]. \tag{5}$$

Using (3)-(5), the cost can be rewritten

$$V = -\bar{\epsilon}\,\mathbb{E}[\tau] + \sum_{k=0}^{\infty}\sum_{i=1}^{n} q_i \mathcal{T}_k(\lambda_i)\left(\epsilon E_L + \eta \mathbb{E}[B_{k+1}]\right). \tag{6}$$

Cost of IEEE 802.16e's sleep policy Our system model enables us to evaluate the cost, denoted V_{Std}, incurred by the sleep policy of the IEEE 802.16e protocol, and more precisely, the sleep policy advocated for type I power saving classes [6]. There, vacations are deterministic (so we use small letters to express that) and the size of a sleep window (i.e., a vacation) is doubled over time until a maximum permissible sleep window, denoted $b_{\max}$, is reached. The size of the kth vacation is then

$$b_k = b_1 2^{\min\{k-1,l\}}, \quad k \in \mathbb{N}^*$$

where $l := \log_2(b_{\max}/b_1)$. We also have

$$t_k = b_1\left(2^{\min\{k,l\}} - 1 + 2^l (k-l)\mathbb{1}\{k > l\}\right), \quad k \in \mathbb{N}^*.$$

The cost of the standard's policy is, using (6),

$$V_{\text{Std}} = -\bar{\epsilon}\,\mathbb{E}[\tau] + \sum_{k=0}^{\infty}\sum_{i=1}^{n} q_i e^{-\lambda_i t_k}\left(\epsilon E_L + \eta b_{k+1}\right), \tag{7}$$

3 Dynamic Programming

Dynamic programming (DP) is a well-known tool which allows to compute the optimal decision policy to be taken at each intermediate observation point, taking into account the whole lifetime of the system. Considering our system model, we want to identify the optimal sleep strategy where decisions are taken at each intermediate wake-up instance. Hence, a DP approach is a natural candidate for determining the optimal policy.

The observation points are at the end of the vacations, i.e., at t_k. The conditional residual off time at a time t is denoted τ_t. We introduce the following DP:

$$V_k^\star(t_k) = \min_{b_{k+1}\geq 0} \Big\{ \mathbb{E}[c(t_k, b_{k+1})] + P(\tau_{t_k} > b_{k+1}) V_{k+1}^\star(t_{k+1}) \Big\}.$$

Here, $V_k^\star(t_k)$ represents the optimal cost at time t_k where the argument t_k denotes the state of the system at time t_k. The terms $P(\tau_{t_k} > b_{k+1})$ and $c(t_k, b_{k+1})$ respectively represent the transition probability and the stage cost at t_k when the control is b_{k+1}. In generic notation, the per stage cost is

$$c(t, b) = \bar{\epsilon}\, \mathbb{E}[(b - \tau_t)1\!\!I\{\tau_t \leq b\}] + \epsilon(E_L + P_S b). \tag{8}$$

We can see that each stage is characterized by the distribution of the residual off time τ_t. The state of the system in sleep mode can then by described by the distribution of τ_t.

In the rest of this section, three cases will be considered following the distribution of the off time. We start with the DP solution for exponential off times, then derive some structural properties of the DP solution for hyper-exponential off times. Last, the case of general off times is considered: structural properties of the optimal policy are found and then suboptimal solutions through DP are discussed.

3.1 Exponential Off Time

When arrivals form a Poisson process with rate λ, both the off time τ and the conditional residual off time τ_t will be exponentially distributed with parameter λ, whatever t is (i.e., whatever stage). The distribution of τ_t is characterized solely by the rate λ. In other words, as time goes on, the state of the system is always represented by the parameter λ. Henceforth, the DP involves a single state, denoted λ.

We are faced with a Markov Decision Process (MDP), a single state λ, a Borel action space (the positive real numbers) and discrete time. Note that the sleep durations are not discrete. However, decisions are taken at discrete embedded times: the kth decision is taken at the end of the $(k-1)$st vacation. Therefore, we are dealing with a discrete time MDP. This is called "negative" dynamic programming [9]. It follows from [4] that we can restrict to stationary policies (that depend only on the state) and that do not require randomization. Since there is only one state (at which decisions are taken) this implies that one can restrict to vacation sizes that have fixed size and that are the same each time a decision has to be taken. In other words, the optimal sleep policy is the constant one. Hence the optimal value is given by the minimization of the following MDP:

$$\begin{aligned} V^\star(\lambda) = \min_{b\geq 0} \big\{ &\bar{\epsilon}\mathbb{E}\big[\big(b - \tau(\lambda)\big)1\!\!I\{\tau(\lambda) \leq b\}\big] \\ &+ \epsilon(E_L + bP_S) + P\big(\tau(\lambda) > b\big)V^\star(\lambda)\big\}. \end{aligned} \tag{9}$$

Proposition 3.1 *The optimal vacation size for exponential off time and the minimal cost are given by*

$$b^\star = -\frac{1}{\lambda}\left(1 + \frac{\lambda\epsilon E_L}{\eta} + W_{-1}\left(-e^{-1-\lambda\epsilon E_L/\eta}\right)\right); \tag{10}$$

$$V^\star(\lambda) = -\frac{1}{\lambda}\left(\bar{\epsilon} + \eta W_{-1}\Big(-e^{-1-\lambda\epsilon E_L/\eta}\Big)\right), \tag{11}$$

where W_{-1} denotes the branch of the Lambert W function[1] *that is real-valued on the interval* $[-\exp(-1), 0]$ *and always below* -1.

3.2 Hyper-Exponential Off Time

We assume in this section that τ is hyper-exponentially distributed with n phases and parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$.

3.2.1 Distribution of the Conditional Residual Off Time τ_t

The tail of τ_t can be computed as follows

$$P(\tau_t > a) = \frac{P(\tau > t + a)}{P(\tau > t)} = \sum_{i=1}^{n} g_i(\mathbf{q}, t)\exp(-\lambda_i a) \tag{12}$$

where

$$g_i(\mathbf{q}, t) := \frac{q_i \exp(-\lambda_i t)}{\sum_{j=1}^{n} q_j \exp(-\lambda_j t)}, \quad i = 1, \ldots, n. \tag{13}$$

We denote $\mathbf{g}(\mathbf{q}, t)$ as the n-tuple of functions $g_i(\mathbf{q}, t)$, $i = 1, \ldots, n$. Observe that $\mathbf{g}(\mathbf{q}, 0) = \mathbf{q}$. The operator $\mathbf{g}$ transforms the distribution $\mathbf{q}$ into another distribution $\mathbf{q}'$ such that $\sum_{j=1}^{n} q'_j = 1$ and $q'_j > 0$.

Equation (12) is nothing but the tail of a hyper-exponential rv having n phases and parameters $\boldsymbol{\lambda}$ and $\mathbf{g}(\mathbf{q}, t)$. Except for the probabilities of the n phases, the off time τ and its residual time τ_t have the same distribution and same parameter $\boldsymbol{\lambda}$. As time goes on, the residual time keeps its distribution but updates its phases' probabilities, through the operator $\mathbf{g}$. It can be shown that

$$g_i(\mathbf{q}, b_1 + b_2) = g_i\big(g_i(\mathbf{q}, b_1), b_2\big). \tag{14}$$

In other words, the operator $\mathbf{g}$ is such that the result of the transformation after $b_1 + b_2$ units of time is the same as that of a first transformation after b_1 units of time, followed by a second transformation after b_2 units of time.

To simplify the notation, we will drop the subscript of the residual off time τ_t, and instead, we will add as argument the current probability distribution (which is transformed over time through the operator $\mathbf{g}$). For instance, if at some point in time, the residual off time has the probability distribution $\mathbf{q}'$, then we will use the notation $\tau(\mathbf{q}')$.

[1] The Lambert W function, satisfies $W(x)\exp(W(x)) = x$. As $y\exp(y) = x$ has an infinite number of solutions y for each (non-zero) value of x, the function $W(x)$ has an infinite number of branches.

The results above can be extended to account for a random passed time T. We have

$$P(\tau > T + a \,|\, \tau > T) = \sum_{i=1}^{n} g_i(\mathbf{q}, T) \exp(-\lambda_i a)$$

where

$$g_i(\mathbf{q}, T) := \frac{q_i \mathcal{T}(\lambda_i)}{\sum_{j=1}^{n} q_j \mathcal{T}(\lambda_j)} = \frac{q_i \mathcal{T}(\lambda_i)}{P(\tau > T)}. \tag{15}$$

There is an abuse of notation in the definition of $g_i(\mathbf{q}, T)$, as this function depends on the *distribution* of T and not on the rv T itself. The function $g_i(\mathbf{q}, T)$ is *not* a rv. Observe that (13), where time is deterministic, is a particular case of (15). Asymptotic properties of $\mathbf{g}$ are provided next.

Define the composition $\mathbf{g}^m(\mathbf{q}, B) = \mathbf{g}\big(\mathbf{g}^{m-1}(\mathbf{q}, B), B\big) = \mathbf{g}(\mathbf{q}, mB)$, where $\mathbf{g}^1(\mathbf{q}, B)$ is the vector whose ith element is given in (15). Assume, without loss of generality, that $\lambda_1 \leq \ldots \leq \lambda_n$. Let $\mathbf{e}(i)$ be the n-dimensional vector whose ith element is 1 and all other elements are zero.

Lemma 3.1 *Fix $\mathbf{q}$ and let $I(\mathbf{q})$ be the smallest j for which $q_j > 0$. The following limit holds (b is te deterministic version of B):*

$$\lim_{m \to \infty} \mathbf{g}^m(\mathbf{q}, b) = \mathbf{e}(I(\mathbf{q})).$$

Lemma 3.1 states that, as time passes, the residual off time's distribution translates its mass towards the phase with the smallest rate, and converges asymptotically irrespective of the initial distribution. This suggests that there exists a threshold on the time after which the optimal policy is the one that corresponds to the optimal policy for state $I(\mathbf{q})$.

Lemma 3.2 *For any $\mathbf{q}$ we have*

$$\lim_{\mathbf{q}' \to \mathbf{q}} V(\mathbf{q}') = V(\mathbf{q}).$$

Lemma 3.2 states that as the state converges, the value also converges to the value at the converged state.

3.2.2 DP Solution

Below we formulate the optimization problem as an MDP where the state space is taken to be the simplex of dimension n, i.e. the set of probability measures over the set $\{1, 2, ..., n\}$. At each stage, the residual off time sees its probability distribution being updated. Let $\mathbf{q}^0$ denote the probability distribution of the *total* off time. It is then the probability distribution of the residual off time at time 0. Thanks to the property (14), the probability distribution of the residual off time at stage $k+1$, i.e., at time t_k, is $\mathbf{q} = \mathbf{g}(\mathbf{q}^0, t_k)$. Henceforth, there is a one to one relation between the stage and the current probability distribution of the residual off time. Without loss of optimality, either of them can be the state in the MDP [2, Sect. 5.4].

The system state is denoted $\mathbf{q}$ and represents the *current* probability distribution of the residual off time. The initial state is $\mathbf{q}^0$. We assume that the controller can choose any time b (a constant or a rv) until he wakes up. The transition probabilities are simply

$$P_{\mathbf{q}, b, \mathbf{q}'} = \mathbb{1}\{\mathbf{q}' = \mathbf{g}(\mathbf{q}, b)\}.$$

We are faced with an MDP with a Borel action space and a state space that is the set of probability vectors $\mathbf{q}$. Note however that, starting from a given $\mathbf{q}$, there is a countable set Q of $\mathbf{q}$'s so that only states within Q can be reached from $\mathbf{q}$. Therefore we may restrict the state space to the countable set Q. We can again use [4] to conclude that we may restrict to policies that choose at each state a non-randomized decision b, and the decision depends only on the current state (and need not depend on the previous history). We next show that there is some $\overline{b}$ such that actions may be restricted to the compact interval $[0, \overline{b}]$ without loss of optimality.

Consider the policy w that takes always a constant one unit length vacation. It is easily seen that the total expected cost, when using policy w, is upper bounded by

$$\overline{v} := \overline{\epsilon} + \epsilon\big(1 + \sup_i 1/\lambda_i\big)(E_L + P_S).$$

Here, $\overline{\epsilon}$ is an upper bound on the expected waiting cost and $1 + \sup_i 1/\lambda_i$ is an upper bound on $\mathbb{E}[X]$, the expected number of vacations, and on $\mathbb{E}[T_X]$, the expected idle time. We conclude that

Lemma 3.3 *For all $\mathbf{q}$, $V(\mathbf{q}) \leq \overline{v}$.*

Lemma 3.4 *Without loss of optimality, one may restrict to policies that take only actions within $[0, \overline{b}]$ where*

$$\overline{b} = \frac{1}{\overline{\epsilon}}\{\overline{v} + 1 + 1/(\min_i \lambda_i)\}.$$

Proof Let u be an ϵ-optimal Markov policy that does not use randomization, where $\epsilon \in (0, 1)$. If $u_i > \overline{b}$ for some i then the expected immediate cost at step i is itself larger than 1 plus the total expected cost that would be incurred under the policy w:

$$\mathbb{E}\big[\big(b - \tau(\mathbf{q})\big)\mathbb{1}\{\tau(\mathbf{q}) \leq b\}\big] > \overline{v} + 1.$$

Thus, by switching from time i onwards to w, the expected cost strictly decreases by at least 1 unit; thus u cannot be ϵ-optimal. ■

We conclude that the MDP can be viewed as one with a countable state space, compact action space, discrete time, and non-negative costs (known as "negative dynamic programming"). Using [9] we then conclude:

(i) The optimal value (minimal cost) is given by the minimal solution of the following DP:

$$V(\mathbf{q}) = \min_{b \geq 0} \big\{\overline{\epsilon}\mathbb{E}\big[\big(b - \tau(\mathbf{q})\big)\mathbb{1}\{\tau(\mathbf{q}) \leq b\}\big] + \epsilon(E_L + bP_S) + P\big(\tau(\mathbf{q}) > b\big)V\big(\mathbf{g}(\mathbf{q}, b)\big)\big\}. \tag{16}$$

(ii) Let $B(\mathbf{q})$ denote the set of all b's that minimize the right hand side of (16) for a given $\mathbf{q}$. Then any policy that chooses at state $\mathbf{q}$ some $b \in B(\mathbf{q})$ is optimal.

The value iteration can be used as an iterative method to compute $V(\mathbf{q})$. Starting with $V_0 = 0$ we write

$$V_{k+1}(\mathbf{q}) = \min_{b \geq 0} \big\{\overline{\epsilon}\mathbb{E}\big[\big(b - \tau(\mathbf{q})\big)\mathbb{1}\{\tau(\mathbf{q}) \leq b\}\big] + \epsilon(E_L + bP_S) + P\big(\tau(\mathbf{q}) > b\big)V_k\big(\mathbf{g}(\mathbf{q}, b)\big)\big\}.$$

Then $V(\mathbf{q}) = \lim_{k\to\infty} V_k(\mathbf{q})$, see [2]. The iteration is to be performed for every possible state $\mathbf{q}$. Lemma 3.1 implies that the moving state, $\mathbf{g}(\mathbf{q}, b)$, converges asymptotically to $\mathbf{e}(I(\mathbf{q}))$. To complete the value iteration, we compute, for a fixed b,

$$\mathbb{E}\big[(b-\tau(\mathbf{q}))\mathbb{1}\{\tau(\mathbf{q}) \le b\}\big] = b - \sum_{i=1}^{n} q_i \frac{1-\exp(-\lambda_i b)}{\lambda_i}.$$

3.3 General Distribution of Off Time

In this section, off times have a general distribution. As a consequence, one can no longer expect that the residual off time will keep the same distribution over time, updating only its parameters. Therefore, the system state is the instant t at which a vacation is to start. We use again τ_t to denote the conditional residual value of τ at time t (i.e., $\tau - t$ given that $\tau > t$.

As a state space, we consider the set of non-negative real numbers. An action b is the duration of the next vacation. We shall assume that b can take value in a finite set. The set of t reachable (with positive probability) by some policy is countable. We can thus assume without loss of generality that the state space is discrete. Then the following holds:

Proposition 3.2

(i) *There exists an optimal deterministic stationary policy.*

(ii) *Let* $V^0 := 0$, $V^{k+1} := \mathcal{L}V^k$, *where*

$$\mathcal{L}V(t) := \min_b \{c(t,b) + P(\tau_t > b)V(t+b)\}$$

where $c(t,b)$ *has been defined in* (8). *Then* V^k *converges monotonically to the optimal value* $V^\star$.

(iii) $V^\star$ *is the smallest nonnegative solution of* $V^\star = \mathcal{L}V^\star$. *A stationary policy that chooses at state* t *an action that achieves the minimum of* $\mathcal{L}V^\star$ *is optimal.*

Proof (i) follows from [9, Thm 7.3.6], and (ii) from [9, Thm 7.3.10]. Part (iii) is due to [9, Thm 7.3.3]. ■

Observe that V^k expresses the optimal cost for the problem of minimizing the total cost over a horizon of k steps.

Proposition 3.3 *Assume that* τ_t *converges in distribution to some limit* $\widehat{\tau}$. *Define* $v(b) := \widehat{c}(b)/[1 - P(\widehat{\tau} > b)]$. *Then*

(i) $\lim_{t\to\infty} V^\star(t) = \min_b v(b)$.

(ii) *Assume that there is a unique* b *that achieves the minimum of* $v(b)$ *and denote it by* $\widehat{b}$. *Then there is some stationary optimal policy* $b(t)$ *such that for all* t *large enough,* $b(t)$ *equals* $\widehat{b}$.

To recapitulate, we have shown, that for a general off time, it is enough to consider deterministic policies to achieve optimal performance. Also, if the residual off time distribution converges in time then the optimal policy converges to the constant policy and in fact becomes constant after finite time (under the appropriate conditions). This can be shown to be the case with the hyper-exponential distribution. Indeed, its residual time converges in distribution to an exponential distribution, having as parameter the smallest among the rates of the hyper-exponential distribution.

3.3.1 Suboptimal policies through dynamic programming

In this section, we propose a suboptimal solution approach using policy iteration for a few stages. For the rest of the stages, we consider a static control that is computed through parametric optimization.

Consider a class of policies in which all vacations are i.i.d. exponentially distributed rvs. We will refer to this class as the "Exponential vacation policy." The optimal total cost under this policy have been found in [1] to depend only on $\mathbb{E}[\tau]$, namely,

$$V_e^\star = \epsilon(P_S\mathbb{E}[\tau] + E_L) + 2\sqrt{\epsilon\eta E_L \mathbb{E}[\tau]}. \tag{17}$$

The optimal control is (cf. [1])

$$b_e^\star = \sqrt{\epsilon E_L \mathbb{E}[\tau])/\eta} = \sqrt{(\epsilon E_L \mathbb{E}[\tau])/(\bar{\epsilon} + \epsilon P_S)}. \tag{18}$$

With one stage policy iteration, the vacations have the form $(b_1, B, B, \ldots)$ where B is an exponentially distributed rv with mean b. We can use DP to compute the optimal policy within this class. The problem is given by

$$V_1^\star(0) = \min_{b\ge 0}\{c(0,b_1) + P(\tau > b_1)V^\star(b_1)\} \tag{19}$$

where $V^\star(b_1)$ is equivalent to $V_e^\star$ in (17) after replacing the off time τ with the residual off time at time b_1, i.e., τ_{b_1}. The optimal control identified through DP is $b_1^\star$ and $b^\star$.

When τ is hyper-exponentially distributed, the system state is the distribution $\mathbf{q}$ which is transformed after each stage through the operator $\mathbf{g}$.

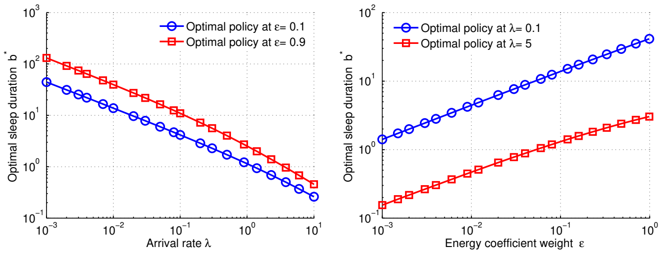
If we add the constraint that the first vacation should be exponentially distributed with the same distribution as B, then we will be back to the problem of finding an optimal exponentially distributed vacation with state-independent mean. Since we do not impose this restriction, the policy obtained after one stage iteration will do strictly better than the Exponential vacation policy.

This suboptimal method for one stage policy iteration can be extended to more stages. Instances of the two stage policy iteration are provided in Sect. 5. As the number of stages of the policy iteration increases, the suboptimal solution converges to the optimal solution (obtained from (16) if τ is hyper-exponentially distributed).

4 Worst Case Performance

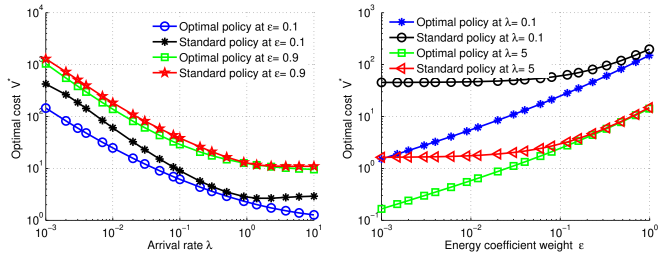
We consider in this section the case where the off time is exponentially distributed with an unknown parameter. When the distribution of the parameter is known (Bayesian framework) the problem reduces to the study of the hyper-exponentially distributed off time. In practice there could be many situations when the statistical distribution of the off time is unknown or hard to estimate. In such non-Bayesian frameworks, we can conduct a worst-case analysis: optimize the performance under the worst case choice of the unknown parameter. We assume that this parameter lies within the interval $[\lambda_a, \lambda_b]$. The worst case is identified as follows

$$\lambda_w := \arg \max_{\lambda\in[\lambda_a,\lambda_b]} \min_{\{B_k\}, k\in\mathbb{N}^*} V \tag{20}$$



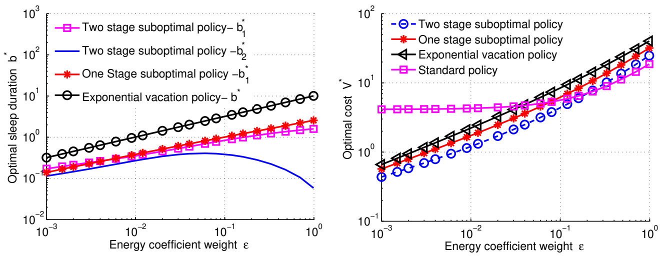
(a) $b^\star$ versus λ (b) $b^\star$ versus ϵ

Figure 1: Optimal sleep duration with exponential off times.



(a) $V^\star$ versus λ (b) $V^\star$ versus ϵ

Figure 2: Optimal expected cost with exponential off times.



(a) Sleep durations versus ϵ (b) Costs versus ϵ

Figure 3: Sleep durations and costs with hyper-exponential off times.

Given that τ is assumed to be exponentially distributed, it is enough to analyze the case of the Constant vacation policy, which has been found to be the optimal in Sect. 3.1. The minimal cost under this policy is given in (11). We have studied (11) using the mathematics software tool, Maple 11. We found the following: $V^\star(\lambda)$ is a monotonic function, decreasing with λ; $\lim_{\lambda\to+\infty} V^\star(\lambda) = \epsilon E_L$; and $\lim_{\lambda\to 0} V^\star(\lambda) = +\infty$. Thus, the optimal control under worst case is the one corresponding to the smallest rate in the interval considered, i.e., $\lambda_w = \lambda_a$.

5 Numerical Investigation

In this section we show some numerical results of our model, when the off time τ is either exponentially or hyper-exponentially distributed. In each case, the best control and the corresponding cost are computed. The cost V_{Std} of the standard's policy is reported (using (7)) for comparison. The physical parameters are set to the following values: $E_L = 10$, and $P_S = 1$. The parameters of the standard protocol are $b_1 = 2$ and $l = 10$.

5.1 Exponential Off Time

In this case, the optimal is to fix all vacations to the value found in (10). This optimal control is depicted in Fig. 1. We naturally find that the optimal sleep duration decreases as λ increases. The physical explanation is that, a large arrival rate forces the server to be available after shorter breaks, otherwise the cost is too high. Also, as ϵ gets smaller, the extra delay gets more penalizing (cf. (2)), enforcing then smaller optimal sleep durations.

Figure 2 depicts the optimal (cf. (11)) and standard (cf. (7)) costs. Observe in Fig. 2(a) how the cost decreases asymptotically to ϵE_L (1 for $\epsilon = 0.1$ and 9 for $\epsilon = 0.9$) as foreseen in Sect. 4. As λ decreases, the increase of the optimal cost is due to the increase of the optimal sleep duration, while for the standard's policy the cost increase is due to the extra (useless and costly) listening. The optimal cost increases with ϵ (cf. Fig. 2(b)). Small values of ϵ make the cost more sensitive to delay, thereby enforcing vacations to be smaller and subsequently incurring smaller costs.

The cost of the standard's policy is high at small ϵ, when delay is very penalizing. This is because the standard has been designed to favor energy over delay. As the vacation size increases exponentially over time, the extra delay can get very large.

5.2 Hyper-Exponential Off Time

In this case, we are able to compute two suboptimal policies using policy iteration. We compare the performance of these to that of the Exponential vacation policy and the standard's policy. The off time distribution is hyper-exponential with parameters $\boldsymbol{\lambda} = \{0.2, 3, 10\}$ and $\mathbf{q} = \{0.1, 0.3, 0.6\}$. The suboptimal solutions are evaluated using (19), the exponential vacation policy using (17)-(18) and the standard's policy using (7).

The performance of the four policies is depicted in Fig. 3 against the energy coefficient weight ϵ. Naturally, the suboptimal policies perform strictly better than the Exponential vacation policy, having the two stage iteration policy strictly outperforming the one stage one (cf. Fig. 3(b)). Interestingly, for large value of ϵ, the standard's policy outperforms all the other policies. As observed earlier, the standard favors energy over delay, so that at large ϵ, it is very efficient in reducing the cost. It is expected however that n-stage policy iteration will outperform the standard for sufficiently large n.

6 Concluding Remarks

We have introduced a model for the control of vacations for optimizing energy saving in wireless networks taking into account the tradeoff between energy consumption and delays. Previous models studied in the literature have considered an exogenous arrival process, whereas we considered an on-off model in which the off duration begins when the server leaves on vacation and where the duration of the on time does not depend on when it starts. We derived the optimal policy in case of a Poisson arrival process and found many structural properties of the optimal policy for hyper-exponential and general off times. Suboptimal policies have been derived in this case using one and two stage policy iteration.

References

[1] A. P. Azad, S. Alouf, E. Altman, V. Borkar, and G. Paschos. Optimal sampling for state change detection with application to the control of sleep mode. Research Report RR-7026, INRIA, Sept. 2009.

[2] D. Bertsekas. *Dynamic Programming and Optimal Control*, volume I. Athena Scientific, second edition, 1996.

[3] M. Crovella and A. Bestavros. Self-similarity in world wide web traffic-evidence and possible causes. In *Proc. of ACM Sigmetrics, Philadelphia, PE*, pages 160–169, 1996.

[4] E. Feinberg. On stationary strategies in borel dynamic programming. *Math. of Operations Research*, 17(2):392–397, May 1992.

[5] A. Feldmann and W. Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance Evaluation*, 31(8):963–976, August 1998.

[6] IEEE Std 802.16e-2005: Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, 2005.

[7] D. G. Jeong and W. S. Jeon. Performance of adaptive sleep period control for wireless communications systems. *IEEE Trans. on Wireless Communications*, 5:3012–3016, November 2006.

[8] N-H. Lee and S. Bahk. MAC sleep mode control considering downlink traffic pattern and mobility. In *Proc. of IEEE VTC 2005-Spring, Stockholm, Sweden*, volume 3, pages 3102–3106, May 2005.

[9] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.

[10] D. Shuman and M. Liu. Optimal sleep scheduling for a wireless sensor network node. In *Proc. of 40th Asilomar Conference on Signals, Systems and Computers (ACSSC)*, pages 1337–1341, Nov. 2006.

[11] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson. Self-similarity through high variability: Statistical analysis of ethernet lan traffic at the source level. In *Proc. of ACM SIGCOMM, Cambridge, MA*, volume 25, pages 110–113, 1995.

[12] J. Xiao, S. Zou, B. Ren, and S. Cheng. An enhanced energy saving mechanism in ieee 802.16e. In *Proc. of IEEE GLOBECOM 2006*, pages 1–5, November 2006.